Delete the paragraph at page 3, lines 14-25 and insert the following paragraph:

In one embodiment, this invention provides an isolated nucleic acid comprising a nucleic acid selected from the group consisting of a nucleic acid encoding any one of Blm open reading frames (ORFs) 8 through 41, and/or a nucleic acid encoding a polypeptide encoded by any one of Blm open reading frames (ORFs) 8 through 41, and/or a nucleic acid amplified by polymerase chain reaction (PCR) using any one of the primer pairs identified in Table II and the nucleic acid of a bleomycin-producing organism as a template. The nucleic acid may comprise one or multiple (*e.g.* two, more preferably 3 or more) bleomycin open reading frames (*i.e. BLM* ORFs 8 through 41). One preferred nucleic acid comprises a nucleic acid encoding a C domain lacking one or more His residues of the conserved HHxxxDG (SEQ ID NO:4) active site for transpeptidation. In another preferred embodiment the nucleic acid comprises a nucleic acid encoding a protein encoded by a gene selected from the group consisting of blmI, blmII, and blmXI.

Delete the paragraphs at page 15, lines 18-31 and insert the following:

Figure 8A shows a restriction map of the *blm* gene cluster from *Sv* ATCC15003 (B, *Bam*HI). 8B shows the relative position of the *blmI*, *blmII*, and *blmXI* genes to the two *blmAB* resistance genes (*blm*^{R}, Blm resistance). Individual open reading frames are represented by open arrows. Figure 8C (SEQ ID NO:127 & 128) shows the nucleotide sequence of the *blmI* gene. The potential ribosome-binding site (RBS) and the conserved motif for 4'-phosphopantetheinylation are underlined. The sequence has been deposited into GenBank under accession no. AF210249.

Figure 9 shows an amino acid sequence comparison of BlmI (SEQ ID NO:133) with PCP domains of known type I NRPSs (Grs-2 [P14688] (SEQ ID NO:129), 36% identity, 58% similarity; Srfa-3 [Q08787] (SEQ ID NO:130), 40% identity, 64% similarity; Vir-s [Y11547] (SEQ ID NO:131), 36% identity, 60% similarity; Saf-b [U24657] (SEQ ID NO:132), 40% identity, 54% similarity). Given in brackets are nucleotide sequence accession numbers. The shaded letters indicate similar amino acids. Consensus residues are amino acids that are similar in more than three sequences. The signature motif for 4'-phosphopantetheinylation is underlined.

Delete the paragraph at page 68, line 8 through page 69, line 16 and insert the following:

The similarities among PPTases from different organisms are reduced to two short motifs separated by 40-45 residues: (V/I)G(V/I)D (SEQ ID NO:87), and (F/W)(S/C/T)XKE(A/S)hhK (SEQ ID NO:91) (Lambalot et al. *Chem. Biol.* (1996) 3:923-936; Walsh etal. *Curr. Opin. Chem. Biol.* (1997)

1:309-315). Our previous attempts to amplify PPTase sequences from *S. verticillus* chromosomal DNA using degenerate primers according to the two conserved motifs were unsuccessful (unpublished results), so we decided to narrow our target. PPTases have been classified in two groups, according to their specificity for the carrier-protein substrate: PPTases involved in polyketide/fatty acid biosynthesis use acyl carrier proteins (ACPs) as substrate, while those for non-ribosomal peptide biosynthesis use peptidyl carrier proteins (PCPs) or aryl carrier proteins (ArCPs) (Walsh et al. *Curr. Opin. Chem. Biol.* (1997) 1:309-315). Several "NRPS-type" PPTase sequences were used to screen the databases to look for actinomycete homologues, and four proteins of unknown function were found: NshC from *Streptomyces actuosus* (Li et al. *Gene* (1990) 91:9-17), SC5A7. 23 from *S. coelicolor* (GenBank AL031107), an unnamed protein from *Streptomyces* sp. strain TH1 (Mori et al. *J. Bacteriol.* (1997) 179:5677-5683), and Rv2794c (later renamed PptT (Quadri et al. *Chem. Biol.* (1998) 5:631-645)) from *Mycobacterium tuberculosis* (GenBank AL008967). The alignment of the actinomycete sequences showed the two motifs conserved in all PPTases and an additional motif - the "THC" motif: PXWPXGX_2GS(M/L)THCXGY (SEQ ID NO:86), located about 15 amino acids upstream of the (V/I)G(V/I)D motif (SEQ ID NO:87). The "THC" motif is not universally conserved in all PPTases, but it can be detected also in some non-actinomycete PPTases like EntD (Coderre et al. *J. Gen. Microbiol.* (1989) 135:3043-3055). Using a recently developed method of PCR primer design (the CODEHOP strategy (COnsensus-DEgenerate Hybrid Oligonucleotide Primer) (Rose et al. *Nucleic Acids Res.* (1998) 26:1628-1635), two primers were designed around the typical C-terminal PPTase motif (primers KEA-1: 5'-T GCA GCA GAA CAG GAG GCK NYC CCA NKG-3' (SEQ ID NO:88) and KEA-2: 5'-TG GGT CAG CGG GTA CCA NRC YTT RWA-3' (SEQ ID NO: 89, H=C+A, N=A+C+T+G, Y=C+T, K=G+T, R=A+G, W=T+A)), and one primer was designed from the "THC" motif (primer THC: 5'-C GGC ATG GTC GGC TCC HTN ACN CAY TG-3', SEQ ID NO:90, H=C+A, N=A+C+T+G, Y=C+T, K=G+T, R=A+G, W=T+A); this motif is not universally conserved in PPTases of all organisms). Using *S. verticillus* chromosomal DNA as template, no amplification product was detected using the THC and the KEA-1 primers. The set of primers THC/KEA-2 successfully amplified a single band of the expected size (about 250 bp), which was gel-purified and cloned. Eight individual clones were sequenced, and all of them resulted to be identical (except differences due to primer utilization) and highly similar to the putative actinomycete PPTases. The PCR fragment was used as a probe to screen a *S. verticillus* genomic library by colony hybridization. Of the 10,000 colonies screened, 25 positive clones were identified, and then confirmed by Southern analysis

a4
cont

to contain the same 4. 6-kb *Bam*HI hybridizing band. The 4. 6-kb DNA fragment was subcloned, and the nucleotide sequence of a 1,761-bp *Bam*HI-*Sal*I region was determined (SEQ ID NO. 3).--

Delete the paragraph on page 69, line 17 through page 70, line 20 and insert the following:

--The sequence of the 1,761-bp *Bam*HI-*Sal*I fragment was analyzed for coding regions by using the CODONPREFERENCE and TESTCODE programs of the GCG package (Genetics Computer Group, Madison, Wisconsin). Two complete ORFs (*pptA, orf3*) and two incomplete ORFs (*orf1, orf4*) were identified within the sequenced region (Figure 13). The first ORF from left to right (designated *orf1*) starts out of the analyzed area and ends with a TGA codon at position 248 of the sequenced fragment. Comparison of the deduced product of *orf1* with proteins in databases showed similarities with Rv2795c from *Mycobacterium tuberculosis* (GenBank AL008967) and SC5A7. 22 from *S. coelicolor* (GenBank AL031107), both of unknown function. The second ORF, *pptA*, contains the sequence amplified by PCR and used for the cloning of this locus. It comprises 741 nucleotides, starting with a GTG codon (position 245) which is coupled to the stop codon of *orf1*, and ending with a TAA codon. The starting codon of *pptA* is preceded by a potential ribosomal binding site (RBS), GGGAG. The overall (76. 6%) and third codon position (93. 9%) G+C contents and the codon usage of *pptA* are similar to those found in other *Streptomyces* genes, with the exception of the stop codon (TAA), which is most uncommon in this group of organisms (Wright et al. *Gene* (1992) 113:55-65). The *pptA* gene encodes a protein of 246 amino acids with a predicted molecular mass of 25,619 Da and a pI of 4. 76, which contains the conserved PPTase motifs. Databases searches with PptA showed significant similarities to the putative actinomycete PPTases (39-52%/48-61% identity/similarity) and to confirmed bacterial PPTases such as EntD from *E. coli* (17%/24% identity/similarity) (Lambalot et al. *Chem. Biol.* (1996) 3:923-936). The third ORF, *orf3*, is separated from *pptA* by an apparently noncoding DNA region of 153 bp, and it is transcribed in opposite and convergent direction with respect to *orf1-pptA*. The gene *orf3* comprises 240 nucleotides, starting with an ATG codon (position 1358) and ending with TGA. The starting codon of *orf3* is preceded by the sequence GAAGG, a potential RBS. The deduced product of *orf3* encodes a protein of 79 amino acids with a predicted mass of 7,555 Da and a pI of 7. 17. The Orf3 protein shows similarities to the N-terminal region of SC5H1. 35c, a protein of unknown function from *S. coelicolor* (GenBank AL049863). Analysis of Orf3 with the SignalP program (Nielsen et al. *Protein Engineer.* (1997) 10:1-6) predicts an N-terminal signal peptide which would be cleaved between residues 27 and 28 (ALA-DS), suggesting that the mature protein (52 amino acids, 5,099 Da, pI 4. 31) would be secreted. Between *orf3* and *orf4* there is an

apparently noncoding region of 251 nucleotides. The *orf4* gene is transcribed in opposite and divergent direction with respect to *orf3*. It starts with an ATG codon at position 1610, preceded by a potential RBS (GGAGG), and ends out of the sequenced fragment. The deduced protein product (50 amino acids) of the incomplete *orf4* contains a potential NAD/FAD binding motif, $GXGX_2GX_3GX_6G$ (SEQ ID NO:92) (Scrutton et al. *Nature* (1990) 343:38-43), showing low similarities to diverse oxidoreductases.--

At pages 21-23, delete Table II and insert the following replacement Table II:

--

Table II. *Blm* gene cluster open reading frames (ORFs) and primers for ORF amplification.

Orf #	Position	Activity	Method	Primers Forward Reverse	Seq ID No.
orf-8 SEQ ID NO:115	76183-77457	Oxygen-independent coproporphyrinogen III oxidase	Gapped-blast comparison[1]	F: ATGAGCCACGCCATCGGA R: TCAGGCGCGTTCGGGGGC	5 6
orf-9 SEQ ID NO:114	74690-76186	ADP-heptose synthase (*blmC*)	Gapped-blast comparison[1]	F: GTGAACACCGACCTGCCC R: TCATGGGGTGTCTCCCTC	7 8
orf-10 SEQ ID NO:113	74421-74693	Peptidyl carrier protein (*blmI*)	Expression and biochemical characterization.[2]	F: ATGAGCGCCCCGCGGGGC R: TCACCGGTCCCGCTCCCC	9 10
orf-11 SEQ ID NO:112	72787-74424	Carbamyltransferase (*blmD*)	Gapped-blast comparison[1]	F: ATGAGCGCCGACCCGTCC R: TCATGAGCGGGCCGCCGT	11 12
orf-12 SEQ ID NO:111	71618-72790	ADP-heptose:LPS heptosyl transferase (*blmE*)	Gapped-blast comparison[1]	F: ATGACCACCCCCATGACC R: TCATGGGGTACTCCTGAT	13 14
orf-13 SEQ ID NO:110	70983-71546	Homolog of mbtH in the synthesis of mycobactin	Gapped-blast comparison[1]	F: ATGACCACGACCCCGCGG R: TCAGGTGCCGGACACGCG	15 16
orf-14 SEQ ID NO:109	69598-70986	Peptide synthetase (condensation, *blmII*)	Gapped-blast comparison[1]	F: GTGACCGCCCCCGGCACA R: TCATCGGTGGCTCCTCGT	17 18
orf-15 SEQ ID NO:108	68582-69601	Regulatory gene (homolog of *syrP*)	Gapped-blast comparison[1]	F: GTGAACCGGCACGGCCCC R: TCACGCGCTCACCTCGTC	19 20
orf-16 SEQ ID NO:107	65778-68585	Mutated peptide synthetase-oxidase (NRPS-0, *blmIII*)	Gapped-blast comparison[1]	F: GTGACGAGCGCCCGGCCC R: TCACGGGGCCTCCGTGCG	21 22
orf-17 SEQ ID NO:106	57901-65781	Peptide synthetase (NRPS-2-1,*blmIV*)	Expression and biochemical characterization.[2]	F: ATGCTGCACGGCGCCGCG R: TCACTCCGGTCCACCTCC	23 24
orf-18 SEQ ID NO:105	55899-57815	Asparagine synthetase	Gapped-blast comparison[1]	F: GTGAGGCCCGTGTGCGGC R: TCAGCCACCGTTGCCGCC	25 26
orf-19 SEQ ID NO:104	54418-55902	Homolog of hydroxylase-dehydrogenase (*blmF*)	Gapped-blast comparison[1]	F: GTGAAGGACCTCGGCCGG R: TCACTCCCCCGGTGCCGG	27 28
orf-20 SEQ ID NO:103	53427-54404	Nucleotide-sugar epimerase (*blmG*)	Gapped-blast comparison[1]	F: GTGACATGGACCGTGGTG R: TCAGGCATCGGCCCTCCC	29 30
orf-21 SEQ ID NO:102	51493-53430	Peptide synthetase (NRPS-3CT, *blmV*)	Gapped-blast comparison[1]	F: ATGCGCGGGCATGACGAC R: TCACGGTGTCTCTCCCTC	31 32
orf-22 SEQ ID NO:101	43263-51290	Peptide synthetase (NRPS-5-4-3, *blmVI*)	Expression and biochemical characterization.[2]	F: ATGAGCCGGCCGGCCGGC R: TCATGCTCGGTCATCGCC	33 34
orf-23 SEQ ID NO:100	39610-43266	Peptide synthetase (NRPS-6, *blmVII*)	Expression and biochemical characterization.[2]	F: GTGACCACGCCCCGCATC R: TCATTCGGGACGCGGGCA	35 36